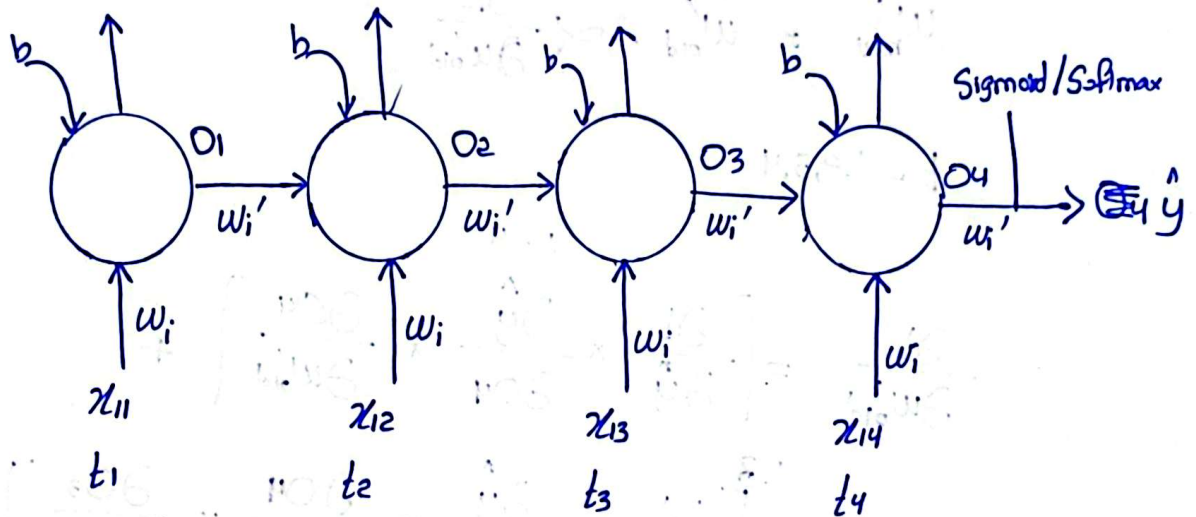


# RNN Back Propagation



## Forward Propagation

$$O_1 = f(x_{11} w_i + b)$$

$$O_2 = f(x_{12} w_i + O_1 w_i' + b)$$

$$O_3 = f(x_{13} w_i + O_2 w_i' + b)$$

$$O_4 = f(x_{14} w_i + O_3 w_i' + b)$$

$$\hat{y} = \sigma(O_4 w_i')$$

$$\text{Loss} = (y - \hat{y})$$

## Backward Propagation

$$w_{\text{new}} = w_{\text{old}} - \alpha \frac{\partial L}{\partial w_{\text{old}}}$$

$$\left[ \frac{\partial L}{\partial w_{\text{old}}} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial w_{\text{old}}} \right] \text{update output weights}$$

### Update hidden layer weights

$$w'_{\text{new}} = w'_{\text{old}} - \alpha \frac{\partial L}{\partial w'_{\text{old}}}$$

$$t = 1, 2, 3, 4$$

$$\begin{aligned} \frac{\partial L}{\partial w'_{\text{old}}} = & \left[ \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial o_4} \times \frac{\partial o_4}{\partial w'_{\text{old}}} \right] + \\ & \left[ \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial o_4} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_3}{\partial w'_{\text{old}}} \right] + \\ & \left[ \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial o_4} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_3}{\partial o_2} \times \frac{\partial o_2}{\partial w'_{\text{old}}} \right] + \\ & \left[ \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial o_4} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_3}{\partial o_2} \times \frac{\partial o_2}{\partial o_1} \times \frac{\partial o_1}{\partial w'_{\text{old}}} \right] \end{aligned}$$

### Update input weights

$$w_{\text{new}} = w_{\text{old}} - \alpha \frac{\partial L}{\partial w_{\text{old}}}$$

$$\begin{aligned} \frac{\partial L}{\partial w_{\text{old}}} = & \left[ \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial o_4} \times \frac{\partial o_4}{\partial w_{\text{old}}} \right] + \left[ \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial o_4} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_3}{\partial w_{\text{old}}} \right] + \\ & \left[ \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial o_4} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_3}{\partial o_2} \times \frac{\partial o_2}{\partial w_{\text{old}}} \right] + \\ & \left[ \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial o_4} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_3}{\partial o_2} \times \frac{\partial o_2}{\partial o_1} \times \frac{\partial o_1}{\partial w_{\text{old}}} \right] \end{aligned}$$